Zero-one laws for binary random fields

David Coupier, Paul Doukhan and Bernard Ycart

(1) MAP5, UMR CNRS 8145, Université René Descartes, Paris

(2) LS-CREST, URA CNRS 220, Paris

(3) LMC-IMAG, UMR CNRS 5523, Grenoble

E-mail address, 1: coupier@math-info.univ-paris5.fr
E-mail address, 2: doukhan@ensae.fr
E-mail address, 3: ycart@imag.fr
URL: http://www-lmc.imag.fr/lmc-sms/Bernard.Ycart/publis/zufields.pdf

Abstract. A set of binary random variables indexed by a lattice torus is considered. Under a mixing hypothesis, the probability of any proposition belonging to the first order logic of colored graphs tends to 0 or 1, as the size of the lattice tends to infinity. For the particular case of the Ising model with bounded pair potential and surface potential tending to $-\infty$, the threshold functions of local propositions are computed, and sufficient conditions for the zero-one law are given.

1. Introduction

There exist essentially three types of zero-one laws. The most ancient (Kolmogorov, Hewitt-Savage) concern asymptotic events defined on a sequence of independent random variables (see Feller (1971) p. 124). More recently, results describing how the probability of an increasing subset of $\{0,1\}^n$ goes from almost 0 to almost 1 have been developed (see for instance Talagrand (1994)). Zero-one laws in logic state that all propositions of a given logic have a probability tending to 0 or 1, as the size of the domain to which they apply tends to infinity. They have a long history: see Compton (1989) or Winkler (1993) for reviews, and Spencer (2001) on logic and random graphs. We shall use Ebbinghaus and Flum (1995) as a basic reference on first order logic.

The first zero-one law in logic was proved independently by Glebskii et al. (1969) and Fagin (1976). It states that the probability of any first order property on a finite domain without constraints tends to 0 or 1 as the size of the domain tends to infinity, under the uniform distribution on the set of structures. It was soon recognized that it also holds when the set of structures is endowed with a product of Bernoulli distributions with fixed parameter p. As an elementary example, consider the model made of a single unary predicate, applied to the domain $V_n = \{0,\ldots,n-1\}$ (see Spencer (2001), Section 0.2). Any structure with k asserted and $n-k$ negated facts has probability $p^k(1-p)^{n-k}$. The Glebskii et al.–Fagin theorem applies, but the expressive power of the language is very poor. Let us add a binary predicate R. The elements of V_n are interpreted as vertices of a graph, and any fact Rxy will be seen as an existing edge between vertices x and y. The unary predicate, denoted by C, will be interpreted as a coloring of vertices: the fact Cx means that x is a black vertex and $\neg Cx$ that it is white. Among all possible 2^{n+n^2} structures, we shall retain only those 2^n for which the graph is cyclic:

$$Rxy \Longleftrightarrow y = x \pm 1 \mod n\ .$$

The expressive power of the language has now increased, since it allows statements about the geometrical arrangement of colored vertices. For instance the first order logic now contains the sentence "there exist 3 black vertices in a row". This model is essentially the "random circular

Received by the editors 5/20/06, revised.
2000 *Mathematics Subject Classification.* 60F20.
Key words and phrases. zero-one law, first-order logic, random field, weak dependence, Ising model.

unary predicate" model of Spencer (1993), Section 2. One can see it as a fixed graph structure (the n-cycle) with randomly colored vertices. The random colors are i.i.d. random variables, Bernoulli with parameter p. The Glebskii et al.–Fagin theorem does not apply anymore, since the set of structures has been restricted. However, it is easy to see that the zero-one law still holds. But until now only *independent* random colors have been considered. The aim of this article is to study situations under which the zero-one law holds even though the colors of vertices are dependent random variables.

Let us generalize the previous model by considering a lattice graph in dimension d, with periodic boundary conditions (lattice torus). The vertex set is now $V_n = \{0, \dots n-1\}^d$. The edge set, denoted by E_n, will be specified by defining the set of neighbors $\mathcal{V}(x)$ of a given vertex x:

$$\mathcal{V}(x) = \{y \neq x \in V_n\,,\ \|y - x\|_p \leq \rho\}\ , \tag{1.1}$$

where the substraction is taken componentwise modulo n, $\|\cdot\|_p$ stands for the L_p norm in $\mathbb{R}^d$ ($1 \leq p \leq \infty$), and ρ is a fixed parameter. For instance, the square lattice is obtained for $p = \rho = 1$. Replacing the L_1 norm by the L_∞ norm adds the diagonals.

A lexical bridge between probability and logic is needed here. A *binary random field* is a set of random variables indexed by V_n, with values in $\{-1, +1\}$ (see Rosenblatt (1985) as a general reference). Obviously the choice $\{-1, +1\}$ for the values of the field is arbitrary: it could be replaced by $\{0, 1\}$ or $\{\text{white}, \text{black}\}$. A binary random field can also be viewed as a random mapping from V_n to $\{-1, +1\}$, which is usually called a *configuration* (a spin configuration for the Ising model). The set of configurations will be denoted by Ξ_n. In the logical interpretation, the graph structure (V_n, E_n) can be described using binary predicates (see Section 2 for details). The values of a configuration can be viewed as a coloring of the graph, and described by a unary predicate C. Let η be a configuration. The assertion $\eta(x) = +1$ will be identified to the fact Cx, and $\eta(x) = -1$ to its negation $\neg Cx$. For convenience, we shall also retain the coloring interpretation:

$$\begin{array}{lcccl} \eta(x) = +1 & \iff & Cx & \iff & \text{vertex } x \text{ is black} \\ \eta(x) = -1 & \iff & \neg Cx & \iff & \text{vertex } x \text{ is white} \end{array}$$

Thus the set of configurations Ξ_n will be identified to the set of those logical structures on the domain V_n for which facts relative to binary predicates are fixed and compatible with the chosen lattice graph. Any closed logical formula (also called sentence) A, expressed in terms of the predicates, determines a subset A_n of Ξ_n that satisfies it. Given a probability distribution μ_n on Ξ_n, the zero-one law holds if $\mu_n(A_n)$ tends to 0 or 1, for each sentence A belonging to the first order logic. Two hypotheses will be used. The first one says that the probability of a local configuration should remain bounded away from 0. The second one is a classical mixing condition: the covariance between functions depending on two disjoint subsets of vertices tends to zero as the distance between the two subsets tends to infinity. Under these hypotheses, Theorem 3.1 shows that the zero-one law holds. The proof is based on Gaifman's theorem (Gaifman (1982) and Ebbinghaus and Flum (1995) p. 31), which says that first-order sentences are essentially local. The zero-one law for first order sentences can be reduced to computing the probability of local patterns (see Section 2 for precise definitions and statements). It will be proved that any local configuration appears with probability tending to 1, an analogue of the "typing monkey" paradox.

Quoting Spencer (1993), "To people who work on Random Graphs the cases p constant are only a small and relatively uninteresting part of the theory". In the circular model above, when $p = p(n)$ depends on the number of vertices, threshold phenomenons similar to those of random graphs occur. For instance, the probability of having k consecutive black vertices tends to 0 if $p(n) \ll n^{-\frac{1}{k}}$, it tends to 1 if $p(n) \gg n^{-\frac{1}{k}}$. An analogous result holds for random colorings of a lattice graph in dimension d: the probability of having k neighboring black vertices tends to 0 or 1 according to whether $p(n)$ is small or large compared to the threshold function $p(n) = n^{-\frac{d}{k}}$ (see Coupier et al. (2004) for an interpretation in the context of random images). In order to extend it to weakly dependent random fields, one needs to choose a parametric family of random field distributions. One of the simplest and most widely studied is the Ising model (see e.g. Georgii (1988); Kindermann and Snell (1980)).

Definition 1.1. *Let $G = (V, E)$ be an undirected graph structure with finite vertex set V and edge set E. Let a and b be two reals. The* Ising model *with parameters a and b is the probability measure $\mu_{a,b}$ on $X = \{-1, +1\}^V$ defined by:*

$$\forall \eta \in X \ , \quad \mu_{a,b}(\eta) = \frac{1}{Z_{a,b}} \exp\left(a \sum_{x \in V} \eta(x) + b \sum_{\{x,y\} \in E} \eta(x)\eta(y) \right) \ ,$$

where the normalizing constant $Z_{a,b}$ is such that $\sum_{\eta \in E} \mu_{a,b}(\eta) = 1$.

In the classical presentation of statistical physics, the elements of X are spin configurations; the definition of $\mu_{a,b}$ involves a temperature parameter which is not relevant here and will be omitted. The parameters a and b are respectively the "surface" and "pair" potentials. For $a > 0$ (respectively $a < 0$), the measure $\mu_{a,b}$ gives a higher weight to those configurations with a large number of $+1$'s (respectively -1's). For $b > 0$, the measure $\mu_{a,b}$ tends to favor groups of neighboring vertices with the same spin, whereas for $b < 0$, any $+1$ will be more likely surrounded by -1's. For $b = 0$, $\mu_{a,0}$ is a product measure: the spins of all vertices are mutually independent, $+1$ with probability $p = e^a/(e^a + e^{-a})$ and -1 with probability $1 - p$. Observe that the model remains unchanged by swapping $+1$ and -1 and replacing a by $-a$. In order to keep a certain coherence with random graphs, we chose to study negative values of a (corresponding to small values of p).

Let us now consider the Ising model $\mu_{a,b}$ on the lattice graph (V_n, E_n) defined above. The potentials $a = a(n)$ and $b = b(n)$ depend on the size of the lattice. The interaction potential $b(n)$ controls the degree of local dependence: we will assume that it remains bounded. Theorem 4.1 expresses the threshold functions of local properties in terms of $e^{a(n)} n^{\frac{d}{2k}}$: if this quantity tends to 0 or $+\infty$ for all k, then the zero-one law holds.

The paper is organized as follows. In Section 2 the logical setting is described, and the zero-one law for first order logic is reduced to the study of basic local and pattern sentences (Proposition 2.5). Weakly dependent random fields will be introduced in Section 3. It will be proved that the probability of any pattern sentence tends to 1, which implies the zero-one law (Theorem 3.1). Section 4 is devoted to the Ising model with varying potentials a and b. A description of threshold functions for basic local sentences will be given, and the zero-one law will be deduced.

2. First order logic

We follow the notations and definitions in Chapter 0 of Ebbinghaus and Flum (1995) for the syntax and semantics of first order logic. The vocabulary contains one unary predicate, denoted by C, and some binary predicates. They apply to the domain $V_n = \{0, \ldots n-1\}^d$. Once the domain and the vocabulary are fixed, the structures are particular models of the predicates, applied to variables in the domain. To any structure, a graph is naturally associated (Ebbinghaus and Flum (1995) p. 26), connecting those pairs of elements $\{x, y\}$ which are such that Rxy or Ryx are satisfied, where R is any of the binary predicates. In each of our structures, we decide to fix the list of facts relative to binary predicates, so that the associated graph is the lattice torus defined by (1.1). It will be denoted by $G_n = (V_n, E_n)$. As usual, the graph distance *dist* is defined as the minimal length of a path between two vertices. The ball of center x and radius r will be denoted by $B(x, r)$.

$$B(x, r) = \{ \, y \in V_n \, ; \, dist(x, y) \leq r \, \} \ .$$

In order to avoid particular cases due to self-overlapping balls, we will always assume that $n > 2\rho r$. If n and n' are both larger than $2\rho r$, the balls $B(x, r)$ in G_n and $G_{n'}$ are isomorphic. Two properties of the balls $B(x, r)$ will be crucial in what follows. The first one is that two balls with same radius are translated of each other:

$$B(x + y, r) = y + B(x, r) \ .$$

(Recall that all operations on vertices are modulo n.) The second one is that for $n > 2\rho r$, the cardinality of $B(x, r)$ depends on r, but neither on x nor on n; it will be denoted by $\beta(r)$.

Observe that whatever the choices of p and ρ in (1.1) the ball $B(x,r)$ is included in the sub-lattice $[x-\rho r, x+\rho r]^d \cap V_n$, and thus

$$\beta(r) \leq (2\rho r + 1)^d \ .$$

In our setting, two structures may differ only by their unary facts. There is a natural bijection between the set of structures and $\{-1,+1\}^{V_n}$: to a structure S corresponds the configuration η defined by:

$$(S \models Cx) \Longleftrightarrow (\eta(x) = +1) \quad ; \quad (S \models \neg Cx) \Longleftrightarrow (\eta(x) = -1) \ .$$

Formulas such as Cx, $Rxy\ldots$ are called *atoms*. The *first-order logic* (Ebbinghaus and Flum (1995) p. 5) is the set of all formulas obtained by recursively combining first-order formulas, starting with atoms.

Definition 2.1. *The set $\mathcal{L}_1$ of first-order formulas is defined by:*

(1) *All atoms belong to $\mathcal{L}_1$.*
(2) *If A and B are first-order formulas, then $(\neg A)$, $(\forall x A x)$ and $(A \wedge B)$ also belong to $\mathcal{L}_1$.*

Observe that the definition of formulas through variables and atoms rules out any reference to a particular vertex. For instance the sentence "the origin is black" does not belong to the first order logic. Moreover, we are interested in formulas for which it can be decided if they are true or false for any given configuration, i.e. for which all variables are quantified. They are called closed formulas, or sentences. Such a sentence A defines a subset A_n of Ξ_n: that of all configurations η that *satisfy* A ($\eta \models A$).

$$A_n = \{\eta \in \Xi_n \, ; \, \eta \models A\} \ .$$

If μ_n denotes a probability measure on Ξ_n, then any sentence A has probability $\mu_n(A_n)$ to be satisfied.

$$\mu_n(\eta \models A) = \mu_n(A_n) \ .$$

Until now, we have not precised which binary predicates belong to the vocabulary. As an example, consider the square torus lattice in dimension 2. Each vertex $x = (i,j)$ has 4 neighbors $(i \pm 1, j), (i, j \pm 1)$ (additions are always taken modulo n). This graph structure can be described by a single binary predicate N (neighbor):

$$Nxy \Longleftrightarrow y = x \pm (0,1) \text{ or } y = x \pm (1,0) \ .$$

One can also consider 2 binary predicates R (right) and U (up).

$$Rxy \Longleftrightarrow y = x + (1,0) \text{ and } Uxy \Longleftrightarrow y = x + (0,1) \ .$$

The underlying graph is the same but the language is more expressive. The sentence "there exist three neighboring black vertices" belongs to both first order logics, but "there exist three horizontally adjacent black vertices" only belongs to the second one. It turns out that the zero-one law depends only on the graph, and not on the binary predicates. This is essentially a consequence of Gaifman's theorem (Ebbinghaus and Flum (1995) p. 31), which states that every first-order sentence is equivalent to a boolean combination of *basic local sentences*.

Definition 2.2. *A basic local sentence has the form:*

$$\exists x_1 \ldots \exists x_m \left(\bigwedge_{1 \leq i < j \leq m} dist(x_i, x_j) > 2r \right) \wedge \left(\bigwedge_{1 \leq i \leq m} \psi_i(x_i) \right) , \tag{2.1}$$

where:

- *m and r are fixed nonnegative integers,*
- *for all $i = 1, \ldots, m$, $\psi_i(x) \in \mathcal{L}_1$ is a formula for which only variable x is free (not bound by a quantifier), and the other variables all belong to the ball $B(x,r)$.*

For fixed radius r, consider a *complete description* $D(0)$ of the ball $B(0,r)$, i.e. a first-order sentence for which all statements concerning vertices at distance at most r of 0 are either asserted or negated. It describes a certain local configuration of the vertices in $B(0,r)$. Denote this

configuration by $\zeta_{D,0}$. For each $x \in V_n$, denote by $\zeta_{D,x}$ the translation of $\zeta_{D,0}$ onto the ball $B(x,r)$:

$$\forall y \in \mathbb{Z}^d\ ,\ dist(0,y) < r \Longrightarrow \zeta_{D,x}(x+y) = \zeta_{D,0}(y)\ .$$

The complete description of the ball $B(x,r)$ that describes the local configuration $\zeta_{D,x}$ will be denoted by $D(x)$.

Definition 2.3. *A single pattern sentence has the form:*

$$\exists x\ D(x)\ , \tag{2.2}$$

where $D(0)$ is a complete description of the ball $B(0,r)$ for a fixed r.

Examples of single pattern sentences are:

(1) "there exists a black vertex",
(2) "there exists a ball of radius 2 with a black vertex at the center, all other vertices in the ball being white".

Definition 2.4. *A pattern sentence has the form:*

$$\exists x_1 \ldots \exists x_m \left(\bigwedge_{1\le i<j\le m} dist(x_i,x_j) > 2r \right) \wedge \left(\bigwedge_{1\le i\le m} D_i(x_i) \right)\ , \tag{2.3}$$

where:

- *m and r are fixed nonnegative integers,*
- *for all $i = 1,\ldots,m$, $D_i(0)$ is a complete description of the ball $B(0,r)$.*

Examples of pattern sentences are:

(1) "there exist 3 black vertices",
(2) "there exist 3 non overlapping white balls of radius 2, one of them with a black vertex at the center".

Obviously, pattern sentences are particular cases of basic local sentences. Proposition 2.5 below reduces the proof of zero-one laws for random configurations to pattern sentences. The set of predicates being fixed, let μ_n be any probability distribution on the set of configurations Ξ_n.

Proposition 2.5. *Consider the following three assertions.*

(i) *The probability of any* pattern *sentence tends to* 0 *or* 1.
(ii) *The probability of any* basic local *sentence tends to* 0 *or* 1.
(iii) *The probability of any* first order *sentence tends to* 0 *or* 1.

Then (i) implies (ii) and (ii) implies (iii).

Proof: Observe first that if the probabilities of sentences A and B tend to 0 or 1, then so do the probabilities of $\neg A$ and $A \wedge B$. This follows from elementary properties of probabilities. As a consequence, if the probability of A tends to 0 or 1 for any A in a given family, this remains true for any finite Boolean combination of sentences in that family. Thus Gaifman's theorem yields that *(ii)* implies *(iii)*. We shall prove now that every basic local sentence is either unsatisfiable or a finite Boolean combination of pattern sentences. Indeed, consider a formula $\psi(x)$ for which only variable x is free, and the other variables all belong to the ball $B(x,r)$. Either it is not satisfiable, or there exists a finite set of local configurations (at most $2^{\beta(r)}$) which satisfy it. To each of those configurations corresponds a complete description $D(x)$ which implies $\psi(x)$. So $\psi(x)$ is equivalent to the disjunction of these $D(x)$'s:

$$\psi(x) \leftrightarrow \bigvee_{D(x)\to\psi(x)} D(x)\ . \tag{2.4}$$

In formula (2.1), one can replace each $\psi_i(x_i)$ by a disjunction of complete descriptions. Rearranging terms, one gets that the basic local sentence (2.1) is itself a finite disjunction of pattern sentences. $\square$

In the case of weakly dependent random fields, to be treated in the next section, we will prove that the probability of any pattern sentence tends to 1. Proposition 2.6 shows that it is enough to examine single pattern sentences. As Proposition 2.5, it holds for any probability distribution μ_n defined on the set of configurations Ξ_n.

Proposition 2.6. *If the probability of any single pattern sentence tends to* 1*, then the probability of any pattern sentence tends to* 1*.*

Proof: Consider the pattern sentence defined by (2.3). We show that its probability tends to 1 by contructing a single pattern sentence that implies it. Consider the ball $B(0,R)$, with $R = m(\rho r{+}1)$. This ball contains m *disjoint* balls of radius r. Denote by $x_1(0),\dots,x_m(0)$ their centers.

The ball $B(0,R)$ is partitioned into $m+1$ sets: the m balls $B(x_1(0),r),\dots,B(x_m(0),r)$ and $\overline{B}(0)$:

$$\overline{B}(0) = B(0,R) \setminus \bigcup_{1\le i\le m} B(x_i(0),r)\ .$$

Define a complete description $\tilde{D}(0)$ on $B(0,R)$ as follows:

(1) For $i = 1,\dots,m$ the restriction of $\tilde{D}(0)$ to $B(x_i(0),r)$ is $D_i(x_i(0))$.
(2) For all $y \in \overline{B}(0)$, $\tilde{D}(0) \models Cy$ (all vertices outside the balls $B(x_i(0),r)$ are black).

Clearly, $\exists x\ \tilde{D}(x)$ implies (2.3). By hypothesis, the probability of $\exists x\ \tilde{D}(x)$ tends to one, hence the result. □

3. Weakly dependent random fields

The notations are those of the previous sections. For each positive integer n, we consider a probability measure μ_n on the set of configurations Ξ_n on the vertices of a lattice graph $G_n = (V_n,E_n)$ in dimension d. Theorem 3.1 below gives conditions under which the probability of any single pattern sentence tends to 1, which implies the zero-one law by Propositions 2.5 and 2.6. The first condition says that the probability of any local configuration remains bounded away from 0. The second one bounds the covariance of two distant configurations. Both conditions, together with the conclusion of the theorem can be expressed in terms of probabilities of events such as "the restriction of the (random) configuration to a given subset coincides with a given local configuration". If η is a (global) configuration on V_n, its restriction to a subset B of V_n will be denoted by η_B. The cardinality of B is denoted by $|B|$. If B and C are two vertex subsets, their distance is defined as usual by:

$$dist(B,C) = \min\{\, dist(x,y)\,,\ x\in B\,,\ y\in C\,\}\ .$$

We call *local event relative to* B an event which depends only on the values of $\eta(x)$ for $x \in B$.

Theorem 3.1. *For each positive integer n, let μ_n be a probability measure on Ξ_n. Assume that the following hypotheses hold.*

(1) *Let r be a fixed integer. For any local configuration $\zeta_{D,0}$ on the ball $B(0,r)$, let $\zeta_{D,x}$ be defined by $\zeta_{D,x}(x+y) = \zeta_{D,0}(y)$, $\forall y \in B(0,r)$. Assume there exists $p_D > 0$ such that*

$$\forall n > 2\rho r\,,\ \forall x\in V_n\,,\quad \mu_n(\eta_{B(x,r)} \equiv \zeta_{D,x}) \ge p_D\ . \tag{3.1}$$

(2) *Assume there exists a function ψ from $\mathbb{N}\times\mathbb{N}$ into $\mathbb{R}$, and a function ϵ from $\mathbb{N}$ to $\mathbb{R}$, decreasing to 0, such that if B, C are disjoint subsets of V_n, and e_B, e_C local events relative to B and C respectively:*

$$|\mu_n(e_B\cap e_C) - \mu_n(e_B)\mu_n(e_C)| \le \psi(|B|,|C|)\,\epsilon(dist(B,C))\ . \tag{3.2}$$

Then for any first order sentence A,

$$\lim_{n\to\infty} \mu_n(A) = 0 \ \textit{or}\ 1\ .$$

Let us notice that the probabilities of local configurations in (3.1) need not be equal: stationarity (spatial homogeneity) is not requested. Observe also that the mixing hypothesis (3.2) is implied by any strong mixing condition such as introduced in Doukhan (1994). After the proof of Theorem 3.1 several examples of random fields satisfying its hypotheses will be given.

Proof: According to Propositions 2.5 and 2.6, it enough to prove that the probability of any pattern sentence tends to 1.

$$\lim_{n\to\infty} \mu_n(\exists x\, D(x)) = 1 \ ,$$

Where $D(0)$ denotes a complete description of the ball $B(0,r)$. Using the notations of the previous section, let $\zeta_{D,0}$ be the unique configuration on $B(0,r)$ described by $D(0)$, and $\zeta_{D,x}$ be its translate on the ball $B(x,r)$:

$$\forall y\,,\ dist(0,y) \leq r\ , \quad \zeta_{D,x}(x+y) = \zeta_{D,0}(y)\ .$$

We need to compute the probability that a random configuration with distribution μ_n coincides with $\zeta_{D,x}$ on $B(x,r)$, for at least one x. We shall first fix a regular sub-lattice of V_n. Let $R > \rho r$ be a given integer, to be specified later. Consider the following set of vertices:

$$T_n = \left\{ \alpha(2R+1)\,,\ \alpha = 0,\ldots,\left\lfloor \frac{n}{2R+1} \right\rfloor - 1 \right\}^d \ , \tag{3.3}$$

where $\lfloor\cdot\rfloor$ denotes the integer part. Denote by τ_n the cardinality of T_n:

$$\tau_n = \left\lfloor \frac{n}{2R+1} \right\rfloor^d \ .$$

Rename the vertices in T_n from 1 to τ_n: $x_1,\ldots,x_{\tau_n}$. Since $R > \rho r$, the balls $B(x_i,r)$ and $B(x_j,r)$ are disjoint for all $i \neq j$. Denote by e_i the event "the configuration coincides with ζ_{D,x_i} on $B(x_i,r)$" and by $\overline{e_i} = \Xi_n \setminus e_i$ the opposite event. Obviously, the union from 1 to τ_n of the e_i's implies $\exists x\, D(x)$. We will prove that the probability of the opposite event tends to 0.

$$\lim_{n\to\infty} \mu_n \left(\bigcap_{1\leq i\leq \tau_n} \overline{e_i} \right) = 0\ .$$

For $i = 1,\ldots,\tau_n$ denote by c_i the following covariance.

$$c_i = \mu_n\left(\bigcap_{j=1}^{i} \overline{e_j}\right) - \mu_n\left(\bigcap_{j=1}^{i-1} \overline{e_j}\right)\mu_n\left(\overline{e_i}\right)\ .$$

An immediate induction leads to:

$$\mu_n\left(\bigcap_{i=1}^{\tau_n}\overline{e_i}\right) = c_{\tau_n} + \sum_{i=1}^{\tau_n-1} c_{n-i} \prod_{l=\tau_n-i}^{\tau_n} \mu_n(\overline{e_l}) + \prod_{l=1}^{\tau_n}\mu_n(\overline{e_l})\ .$$

Therefore:

$$\mu_n\left(\bigcap_{i=1}^{\tau_n}\overline{e_i}\right) \leq \sum_{i=2}^{\tau_n}|c_i| + \prod_{i=1}^{\tau_n}\mu_n(\overline{e_i})\ .$$

By the first hypothesis of the theorem, the product can be bounded as follows.

$$\prod_{l=1}^{\tau_n}\mu_n(\overline{e_l}) \leq (1-p_D)^{\tau_n}\ .$$

In order to make sure that it tends to zero, it is enough to require that τ_n should tend to infinity. Let us now turn to the sum of covariances, for which the second hypothesis has to be used. Observe that by construction of T_n, the distance between any two balls $B(x_i,r)$ and $B(x_j,r)$ is larger than $2(R-\rho r)/\rho$. Thus the second hypothesis implies that for $i = 1,\ldots,\tau_n$,

$$\sum_{i=2}^{\tau_n}|c_i| \leq \sum_{i=2}^{\tau_n}\psi(i\beta(r),\beta(r))\,\epsilon(2(R-\rho r)/\rho)\ .$$

Let us choose a be a positive increasing real function Ψ, tending to infinity, such that for all $m \in \mathbb{N}$:

$$\sum_{i=2}^{m} \psi(i\beta(r), \beta(r)) \leq \Psi(m) \ .$$

Denote by Ψ^{-1} its reciprocal function, which is also positive, increasing, and tends to infinity. Define $R = R_n$ as follows.

$$R_n = \frac{n}{2((\Psi^{-1}(\epsilon(\sqrt{n})^{-1/2}))^{1/d} - 1)} \ .$$

One has :

$$\tau_n = \left\lfloor \frac{n}{2R_n + 1} \right\rfloor^d = \left\lfloor \frac{1}{\frac{1}{n} + \frac{1}{\Psi^{-1}(\epsilon(\sqrt{n})^{-1/2}))^{1/d} - 1}} \right\rfloor^d \ ,$$

which tends to $+\infty$, since ϵ tends to zero, and Ψ^{-1} to $+\infty$. Now :

$$\Psi(\tau_n) \leq \Psi\left(\left(\frac{n}{2R_n} + 1\right)^d\right) = \Psi(\Psi^{-1}(\epsilon(\sqrt{n})^{-1/2})) = \epsilon(\sqrt{n})^{-1/2} \ .$$

Replacing Ψ by a larger function if necessary, one can make sure that $2(R_n - \rho r)/\rho \geq \sqrt{n}$. Since ϵ is decreasing,

$$\epsilon(2(R_n - \rho r)/\rho) \leq \epsilon(\sqrt{n}) \ ,$$

hence:

$$\Psi(\tau_n)\, \epsilon(2(R_n - \rho r)/\rho) \leq \epsilon(\sqrt{n})^{1/2} \ ,$$

which tends to 0 as desired. □

Classical examples of random fields are stationary (i.e. translation invariant) random fields, defined on $\mathbb{Z}^d$ as sets of random variables $\{\eta(x)\,,\ x \in \mathbb{Z}^d\}$ (see Rosenblatt (1985) or Doukhan (1994)). In what follows, we shall view our measure μ_n as the distribution of the restriction $\{\eta(x)\,,\ x \in V_n\}$. Observe that if the distribution of a stationary random field $\{\eta(x)\,,\ x \in \mathbb{Z}^d\}$ satisfies the hypotheses of Theorem 3.1, for $B, C \subset \mathbb{Z}^d$, then the same holds for $B, C \in V_n$ because distances in the torus V_n are always smaller than in $\mathbb{Z}^d$. Notice also, that even if $\{\eta(x)\,,\ x \in \mathbb{Z}^d\}$ is stationary, its restriction to V_n is not stationary in the sense of translations on the torus.

- **Gaussian models**
  They are often used in practice, for instance to model the sea level. Let $Y = \{Y(x)\,,\ x \in \mathbb{Z}^d\}$ be a stationary Gaussian random field, and $h : \mathbb{R} \to \{-1, +1\}$ a discrete valued function. Let μ be the distribution of $\{h(Y(x))\,,\ x \in \mathbb{Z}^d\}$. Assume that the spectral density f_Y of Y is bounded away from zero. Then in (3.2), one can choose $\psi \equiv 1$ as a constant and $\epsilon(s)$ as the distance from f_Y to the set of trigonometric polynomials with degree $< s$.
- **Ising models**
  Another classical example of weakly dependent random field is the Ising model of Definition 1.1. In their famous article Dobrushin and Shlosman (1985) proved the equivalence of several mixing conditions, including the exponential decay of correlations. Those conditions are satisfied in the high temperature case (i.e. for $|\,b\,|$ small enough) or in large external field (i.e. for $|\,a\,|$ sufficiently large). See also Higuchi (1993) for the 2-dimensional case. Theorem 3.1 thus yields the zero-one law for the Ising model with fixed parameters. In Section 4, the parameters a and b will be allowed to depend on n.
- **Weak dependence and mixing**
  The following setting is described in Doukhan and Louhichi (1999) and we shall keep their notations. Let B be a finite subset of $\mathbb{Z}^d$. Consider the set $\Lambda^{(1)}$ of Lipschitz functions $\mathbb{R} \to \mathbb{R}$ such that

$$\mathrm{Lip} f = \sup_{x,y \in \{0,1\}^B} \frac{|f(x) - f(y)|}{\sum_{i \in B} |x_i - y_i|} \leq 1, \quad \|f\|_\infty \leq 1 \ .$$

The so-called $(\epsilon, \Lambda^{(1)}, \phi)$-weak dependence condition of Doukhan and Louhichi (1999), defined from a decreasing sequence $\epsilon_r \searrow 0$ and a function $\phi : \mathbb{N}^2 \times \mathbb{R}^{+2}) \to \mathbb{R}^+$, holds if

$$|\mathrm{Cov}(f_1((X_i)_{i\in B_1}), f_2((X_i)_{i\in B_2})| \le \phi(|B_1|, |B_2|, \mathrm{Lip} f_1, \mathrm{Lip} f_2)\, \epsilon_r \tag{3.4}$$

if $f_1, f_2 \in \Lambda^{(1)}$ are as above with $B_1, B_2 \subset \mathbb{Z}^d$ finite and distant by at least r. A random field satisfying relation (3.4) for the function $\phi(u, v, U, V) = uU + vV$ is called an η-weakly dependent random field.

In our setting, the X_t's take their values in $\{-1, +1\}$. In this case, indicators of local configurations may be written as functions $f((X_i)_{i\in B})$ for a bounded set $B \subset \mathbb{Z}^d$. Indeed, such functions $f : \{-1, 1\}^B \to \{0, 1\}$ are in $\Lambda^{(1)}$ and (3.2) holds with $\psi(u, v) = \phi(u, v, 1, 1)$ and ϵ_r is as in equation (3.4).

So in the binary case the weak dependence condition of Doukhan and Louhichi (1999) implies the mixing condition (3.2). In more general settings, Doukhan and Louhichi (1999) provides a weaker alternative to mixing conditions.

Here are some examples of $(\epsilon, \Lambda^{(1)}, \phi)$-weakly dependent fields.

Bernoulli shifts are defined through an independent and identically distributed random field $(\xi_t)_{t\in\mathbb{Z}^d}$ by the relation $X_n = H((\xi_{n-t})_{t\in\mathbb{Z}^d})$. If such expressions are (a.s.) well defined, they generate stationary random fields. If

$$\delta_r = \mathbb{E}\left|H((\xi_t)_{t\in\mathbb{Z}^d}) - H((\xi_t \mathbf{1}_{(|t|\le r)})_{t\in\mathbb{Z}^d})\right|$$

where $|\cdot|$ denotes a norm on $\mathbb{Z}^d$, then it is easy to prove with Doukhan and Louhichi ideas in Doukhan and Louhichi (1999) that (3.4) holds with $\phi(u, v, U, V) = uU + vV$ and $\epsilon(r) = 2\delta_{r/2}$. In the case where H takes values in $\{-1, +1\}$, one has $\psi(u, v) = u + v$.

Other models have similar expressions, in particular mappings $X_t = h(Y_t)$ of general Bernoulli shifts (Y_t) through functions with values in $\{-1, +1\}$. As a simple example, consider a linear random field, defined by:

$$Y_t = \sum_{s\in\mathbb{Z}^d} a_s \xi_{t-s},$$

where $\delta_r \le \mathbb{E}|\xi_0| \sum_{|i|\ge s} |a_i|$. More general polynomials may also be considered, they are usually named as Chaotic Volterra expansions (see Doukhan and Lang (2002)).

The next example provides a way to exhibit such expansions for $\mathbb{R}^k$-valued random fields. It comes from the finance literature where heteroskedasdicity is a natural feature. Let $(\xi_t)_{t\in\mathbb{Z}^d}$ be a stationary sequence of random $k \times m$ matrices, $(a_j)_{j\in\mathbb{Z}^*}$ be a sequence of $m \times k$ matrices, and a be a vector in $\mathbb{R}^m$. A vector valued $ARCH(\infty)$ random field model is a solution of the recurrence equation

$$X_t = \xi_t \left(a + \sum_{j\ne 0} a_j X_{t-j}\right),\ t \in \mathbb{Z}^d \tag{3.5}$$

Such $ARCH(\infty)$ models include a large variety of models (see Doukhan et al. (2006)). Assume that $\sum_{j\ne 0} \|a_j\| \|\xi_0\|_\infty < 1$, then a stationary solution of equation (3.5) in L^p is given by :

$$X_t = \xi_t \left(a + \sum_{k=1}^{\infty} \sum_{j_1,\ldots,j_k\ne 0} a_{j_1}\xi_{t-j_1} a_{j_2} \ldots a_{j_k} \xi_{t-j_1-\cdots-j_k} a\right) \quad (\in \mathbb{R}^k)\ . \tag{3.6}$$

Denote now $A = \sum_{j\ne 0} \|a_j\|$, $A(x) = \sum_{\|j\|\geqslant x} \|a_j\|$ and $\lambda = \left(\sum_{j\ne 0}^{\infty} \|a_j\|\right) \|\xi_0\|_\infty$, where

$$\|(j_1, \ldots, j_k)\| = |j_1| + \cdots + |j_k|\ .$$

In what follows, we assume for simplicity that the random field $(\xi_t)_{t\in\mathbb{Z}^d}$ is iid.

One approximates here X_t by a random variable independent of X_0. Set

$$\tilde{X}_t = \xi_t \left(a + \sum_{k=1}^{\infty} \sum_{\|j_1\|+\cdots+\|j_k\|<t} a_{j_1}\xi_{t-j_1} \cdots a_{j_k}\xi_{t-j_1-\cdots-j_k} a \right)$$

A bound for the error is given by :

$$\mathbb{E}\|X_t - \tilde{X}_t\| \leqslant \mathbb{E}\|\xi_0\| \left(\mathbb{E}\|\xi_0\| \sum_{k=1}^{t-1} k\lambda^{k-1} A\left(\frac{t}{k}\right) + \frac{\lambda^t}{1-\lambda} \right) \|a\|$$

For suitable constants K and K' we derive

$$\|X_t - \tilde{X}_t\| \leqslant \begin{cases} K\frac{(\log(t))^{\max\{b,1\}}}{t^b} \text{ if } A(x) \leqslant Cx^{-b},\ b, C > 0 \\ K'(\max\{q,\lambda\})^{\sqrt{t}} \text{ if } A(x) \leqslant Cq^x \end{cases}$$

Weak dependence of the corresponding random fields follows. The solution (3.6) of equation (3.5) is $\eta-$weakly dependent with

$$\epsilon_r = \|\xi_0\|_\infty \|a\| x_r, \text{ with } x_t = \|\xi_0\|_\infty \sum_{k=1}^{t-1} k\lambda^{k-1} A\left(\frac{t}{k}\right) + \frac{\lambda^t}{1-\lambda}.$$

More precisely, for Riemannian decays of $\|a_j\|$, with $A(x) \leqslant Cx^{-b}$ for $b, C > 0$, we have

$$x_t \leqslant \begin{cases} \mathbb{E}\|\xi_0\| \left(2\mathbb{E}\|\xi_0\| C\left(\frac{-b}{\ln\lambda}\right)^b \left(\frac{\ln(t)}{t}\right)^b + \frac{\lambda^t}{1-\lambda} \right) \|a\| \text{ if } b > 1 \\ \mathbb{E}\|\xi_0\| \left(3AC\left(\frac{-b}{\ln\lambda}\right) \frac{\ln(t)}{t^b} + \frac{\lambda^t}{1-\lambda} \right) \|a\| \text{ if } 0 < b < 1 \end{cases}$$

For geometric decays $\|a_j\| \leq q^j$,

$$x_t \leqslant \mathbb{E}\|\xi_0\| \left(\sum_{k=1}^{t-1} \mathbb{E}\|\xi_0\|^k \sum_{p=t}^{\infty} C_p^k q^p + \frac{\lambda^t}{1-\lambda} \right) \|a\|$$

In those cases, there exist positive constants K, b, and C such that

$$x_t \leqslant \begin{cases} K\frac{(\log(t))^{\max\{b,1\}}}{t^b}, \text{ under Riemannian decay } A(x) \leqslant Cx^{-b} \\ K(\max\{q,\lambda\})^{\sqrt{t}}, \text{ under geometric decay } A(x) \leqslant Cq^x \end{cases}$$

Now any function $h(X_t)$ of such a random field through a mapping with values in $\{-1,+1\}$ satisfies the assumptions of Theorem 3.1.

Let us finally mention the case of an *associated random field*, i.e. a $\{-1,+1\}$-valued random field η on V_n such that for each coordinatewise non-decreasing function $f, g : \mathbb{R}^{V_n} \to \mathbb{R}$ such that $\mathbb{E}(f^2(X) + g^2(X)) < \infty$, the covariance between $f(\eta)$ and $g(\eta)$ is non negative (this is also known as the FKG condition). Such a field can be treated as before with $\psi(u,v) = uv$ and $\epsilon(s) = \sup_{|u-v|\geq s} \mathrm{Cov}(X_u, X_v)$.

4. Ising model

Here we consider the particular case where the measure μ_n on the set of configurations Ξ_n is the Ising model of Definition 1.1:

$$\forall \eta \in \Xi_n\ , \quad \mu_n(\eta) = \frac{1}{Z} \exp\left(a(n) \sum_{x\in V_n} \eta(x) + b(n) \sum_{\{x,y\}\in E_n} \eta(x)\eta(y) \right)\ .$$

Expectations relative to μ_n will be denoted by $\mathbb{E}_n$. The surface potential $a = a(n)$ and the pair potential $b = b(n)$ both depend on n. We will assume that $a(n)$ is negative: we already observed that the model remains unchanged through swapping -1 and $+1$ and changing a to $-a$. If the pair potential b remains bounded, then conditions on $a(n)$ can be given for the zero-one law.

Theorem 4.1. *Assume that for all n, $a(n) < 0$ and there exists $b_0 > 0$ such that $|b(n)| < b_0$. If*

$$\forall k = 1, 2, \ldots \, , \quad \lim_{n\to\infty} e^{a(n)} n^{\frac{d}{2k}} \in \{0\, , +\infty\} \, ,$$

then for any first order sentence A

$$\lim_{n\to\infty} \mu_n(A_n) = 0 \text{ or } 1 \, .$$

As a particular case, for $b(n) \equiv 0$, the spins are independent, $+1$ with probability $p(n) = e^{a(n)}/(e^{a(n)} + e^{-a(n)}) \sim e^{2a(n)}$ or -1 with probability $1 - p(n)$. The threshold functions for $p(n)$ are $n^{-\frac{d}{k}}$, $k = 1, 2 \ldots$, as expected. By Proposition 2.5, it would suffice to prove that the functions $n^{-\frac{d}{2k}}$ are the threshold functions (for $e^{a(n)}$) of pattern sentences. Actually, it is possible to characterize the threshold functions of all basic local sentences. We begin with the particular case of single pattern sentences.

Proposition 4.2. *Assume that for all n, $a(n) < 0$ and there exists $b_0 > 0$ such that $|b(n)| < b_0$. Let r and k be two integers such that $0 \le k \le \beta(r)$. Let $D(0)$ be a complete description of the ball $B(0, r)$, with exactly k pluses.*

$$\text{If } \lim_{n\to\infty} e^{2a(n)k} n^d = 0 \text{ then } \lim_{n\to\infty} \mu_n(\exists x\ D(x)) = 0 \, . \tag{4.1}$$

$$\text{If } \lim_{n\to\infty} e^{2a(n)k} n^d = +\infty \text{ then } \lim_{n\to\infty} \mu_n(\exists x\ D(x)) = 1 \, . \tag{4.2}$$

If $a(n)$ is such that $e^{a(n)} \sim cn^{-\frac{d}{2k}}$, Coupier (2006) proves that a Poisson approximation holds for "$\exists x\ D(x)$" if $D(0)$ has k pluses. Thus $\mu_n(\exists x\ D(x))$ converges to a limit which is different from 0 and 1.

Proof: Let us start with some notations and definitions. Given $\eta \in E_n = \{-1, +1\}^{V_n}$ and $U \subset V_n$, we denote by η_U the natural projection over $\{-1, +1\}^U$. If U and V are two disjoint subsets of V_n then $\eta_U \zeta_V$ is the configuration on $U \cup V$ which is equal to η on U and ζ on V. We denote by δU the neighborhood of U (corresponding to (1.1)):

$$\delta U = \{y \in V_n \setminus U,\ \exists x \in U,\ \{x, y\} \in A_n\} \, ,$$

and by $\overline{U}$ the union of the two disjoint sets U and δU. In the case of balls, $\overline{B(x, r)} = B(x, r + 1)$. Finally, $|U|$ denotes the cardinality of U and $\mathcal{F}(U)$ the σ-algebra generated by the configurations of $\{-1, +1\}^U$.

Let us fix a complete description $D(0)$ of the ball $B(0, r)$, and denote by $\zeta_{D,x}$ the local configuration on $B(x, r)$ defined by the translation of $D(0)$ (notations of Theorem 3.1). For any vertex x, we define the indicator function I_x^D as follows:

$$\forall \eta \in \Xi_n,\ I_x^D(\eta) = \begin{cases} 1, & \text{if } \eta_{B(x,r)} = \zeta_{D,x} \, , \\ 0, & \text{otherwise} \, . \end{cases}$$

The random variables I_x^D indicate occurrences of a given local configuration on the lattice. Their sum counts these occurrences. It will be denoted by X_n^D.

$$X_n^D = \sum_{x \in V_n} I_x^D \, .$$

Due to periodicity, this sum bears over $|V_n| = n^d$ terms. The event $X_n^D > 0$ is logically equivalent to the single pattern sentence $\exists x\ D(x)$. So, the quantity we are actually interested in is $\mu_n(X_n^D > 0)$.

Let us now fix a vertex x and simply denote by B the ball $B(x, r)$. Recall that there exists a finite number of configurations on B (precisely $2^{\beta(r)}$). Denote by $\mathcal{D}_r^x$ their set. The *local energy* $H^B(\sigma)$ of the configuration $\sigma \in \{-1, +1\}^{\overline{B}}$ on $\overline{B}$ is defined by:

$$H^B(\sigma) = a(n) \sum_{y \in B} \sigma(y) + b(n) \sum_{\substack{\{y,z\} \in A_n \\ (y \in B) \vee (z \in B)}} \sigma(y)\sigma(z) \, . \tag{4.3}$$

Hence, for all $\sigma \in \{-1,+1\}^{\delta B}$, the local energy $H^B(\zeta_{D,x}\sigma)$ on B of the configuration which is equal to $\zeta_{D,x}$ on B and σ on δB, can be expressed as:

$$H^B(\zeta_{D,x}\sigma) = a(n)(2k-\beta(r)) + b\left(\sum_{\substack{\{y,z\}\in A_n\\ y,z\in B}} \zeta_{D,x}(y)\zeta_{D,x}(z) + \sum_{\substack{\{y,z\}\in A_n\\ y\in B, z\in\delta B}} \zeta_{D,x}(y)\sigma(z)\right) ,$$

where k is the number of $+1$'s of the complete description $D(0)$. This notion of local energy naturally appears in the expression of the conditional probability $\mu_n(I_x^D = 1\,|\,\sigma)$, $\sigma \in \{-1,+1\}^{\delta B}$:

$$\mu_n(I_x^D = 1 \mid \sigma) = \frac{e^{H^B(\zeta_{D,x}\sigma)}}{\sum_{\zeta'\in\mathcal{D}_r^x} e^{H^B(\zeta'\sigma)}} . \tag{4.4}$$

Moreover, it is connected to the number X_n^D of copies of $\zeta_{D,0}$ in the graph since:

$$\mathbb{E}_n[X_n^D] = \mathbb{E}_n\left[n^d \mu_n(I_x^D = 1 \mid \mathcal{F}(\delta B))\right] .$$

Here, $\mu_n(I_x^D = 1 \mid \mathcal{F}(\delta B))$ represents a $\mathcal{F}(\delta B)$-measurable random variable, whereas for $\sigma \in \{-1,+1\}^{\delta B}$, $\mu_n(I_x^D = 1\,|\,\mathcal{F}(\delta B))(\sigma) = \mu_n(I_x^D = 1\,|\,\sigma)$ is a conditional probability.

Note that the set δB has a bounded cardinality. Thus, the pair potential $b(n)$ being bounded, there exist two constants m and M (which do not depend on n) such that for all local configuration $\zeta' \in \mathcal{D}_r^x$ having exactly k' pluses and for all configuration $\sigma \in \Xi_n$:

$$2a(n)(k'-k) + m \leq H^B(\zeta'\sigma_{\delta B}) - H^B(\zeta_{D,x}\sigma_{\delta B}) \leq 2a(n)(k'-k) + M . \tag{4.5}$$

Statements (4.1) and (4.2) rely on the fact that the difference between the local energies $H^B(\zeta'\sigma_{\delta B})$ and $H^B(\zeta_{D,x}\sigma_{\delta B})$ is uniformly bounded, for any configuration on δB.

Assume that $k > 0$ and $\lim e^{2a(n)k} n^d = 0$. Let us denote by ζ^0 the unique element of $\mathcal{D}_r^x$ having no $+1$. Relations (4.4) and (4.5) imply that, for all $\sigma \in \Xi_n$,

$$\begin{aligned} \mu_n(I_x^D = 1 \mid \sigma_{\delta B}) &\leq e^{H^B(\zeta_{D,x}\sigma_{\delta B}) - H^B(\zeta^0\sigma_{\delta B})} \\ &\leq e^{2a(n)k-m} . \end{aligned}$$

As a consequence:

$$\begin{aligned} \mathbb{E}_n[X_n^D] &= n^d \mathbb{E}_n\left[\mu_n(I_x^D = 1 \mid \mathcal{F}(\delta B))\right] \\ &\leq n^d e^{2a(n)k-m} , \end{aligned}$$

which tends to 0 as n tends to infinity. Finally, X_n^D being an integer valued variable, its probability of being positive is bounded by its expectation and relation (4.1) follows.

Assume now that $\lim e^{2a(n)k} n^d = +\infty$. For all local configurations $\zeta' \in \mathcal{D}_r^x$ and for all $\sigma \in \Xi_n$, (4.5) implies that the difference $H^B(\zeta'\sigma_{\delta B}) - H^B(\zeta_{D,x}\sigma_{\delta B})$ is bounded by $-2a(n)k + M$ (the surface potential $a(n)$ is negative). Hence, using relation (4.4), there exists a positive constant $c = 2^{-\beta(r)} e^{-M}$ satisfying:

$$\forall \sigma \in \Xi_n,\ \mu_n(I_x^D = 1 \mid \sigma_{\delta B}) \geq c e^{2a(n)k} . \tag{4.6}$$

As in the proof of Theorem 3.1, let us introduce the subset T_n of V_n defined by:

$$T_n = \left\{ \alpha(\rho(2r+1)+1),\ \alpha = 0, \ldots, \left\lfloor \frac{n}{\rho(2r+1)+1} \right\rfloor - 1 \right\}^d ,$$

Denote by τ_n the cardinality of T_n and $\mathcal{T}_n$ the union of the balls of radius r, centered at the elements of T_n (see Figure 4). Firstly, let us remark that there exists a positive constant c' satisfying $\tau_n \geq c' n^d$. Secondly, if x and x' are two distinct elements of T_n then no vertex of the ball $B(x,r)$ can be neighbor with a vertex of the ball $B(x',r)$. Thus, let us define the random variable $\tilde{X}_n^D$ by:

$$\tilde{X}_n^D = \sum_{x\in T_n} I_x^D .$$

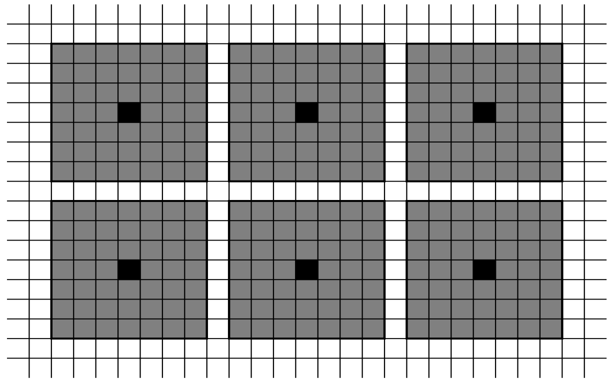

FIGURE 4. The grey vertices represent the set $\mathcal{T}_n$ (in dimension $d=2$, with $r=3$, $\rho=1$, and $p=+\infty$). The black vertices are elements of T_n.

As $\tilde{X}_n^D \leq X_n^D$, it suffices to prove that $\mu_n(\tilde{X}_n^D = 0)$ tends to 0 (to obtain (4.2)). The Gibbs measure μ_n is a Markov random field with respect to the graph structure defined by (1.1). This has two consequences. On the one hand, due to the definition of T_n, one has:

$$\begin{aligned} \mu_n\left(\tilde{X}_n^D = 0 \mid \mathcal{F}(\delta\mathcal{T}_n)\right) &= \mu_n\left(\bigwedge_{x\in T_n} I_x^D = 0 \mid \mathcal{F}(\delta\mathcal{T}_n)\right) \\ &= \prod_{x\in T_n} \mu_n\left(I_x^D = 0 \mid \mathcal{F}(\delta\mathcal{T}_n)\right) . \end{aligned} \tag{4.7}$$

On the other hand, for all $x \in T_n$, the random variable $\mu_n(I_x^D = 0 \mid \mathcal{F}(\delta\mathcal{T}_n))$ is $\mathcal{F}(\delta B(x,r))$-measurable:

$$\mu_n\left(I_x^D = 0 \mid \mathcal{F}(\delta\mathcal{T}_n)\right) = \mu_n\left(I_x^D = 0 \mid \mathcal{F}(\delta B(x,r))\right) . \tag{4.8}$$

Let σ^{max} be an element of Ξ_n maximizing the conditional probability $\mu_n(I_x^D = 0 \,|\, \sigma_{\delta B(x,r)})$. Then, from (4.7) and (4.8), we get for any vertex x and for any $\sigma \in \Xi_n$:

$$\begin{aligned} \mu_n\left(\tilde{X}_n^D = 0 \mid \mathcal{F}(\delta\mathcal{T}_n)\right) &\leq \left(\mu_n(I_x^D = 0 \mid \sigma^{max}_{\delta B(x,r)})\right)^{\tau_n} \\ &\leq e^{-\tau_n \mu_n\left(I_x^D = 1 | \sigma^{max}_{\delta B(x,r)}\right)} . \end{aligned}$$

Finally, the inequality (4.6) implies that:

$$\begin{aligned} \mu_n\left(\tilde{X}_n^D = 0\right) &= \mathbb{E}_n\left[\mu_n\left(\tilde{X}_n^D = 0 \mid \mathcal{F}(\delta\mathcal{T}_n)\right)\right] \\ &\leq e^{-\tau_n \mu_n\left(I_x^D = 1 | \sigma^{max}_{\delta B(x,r)}\right)} \\ &\leq e^{-c\, c' n^d e^{2a(n)k}} \end{aligned}$$

which tends to 0 as n tends to infinity by hypothesis. □

Proposition 4.2 shows that the appearance of a given local pattern only depends on its number of +1's: if $e^{a(n)}$ is small compared to $n^{-\frac{d}{2k}}$, then no pattern of fixed size, with k pluses, should appear in the configuration. If $e^{a(n)}$ is large compared to $n^{-\frac{d}{2k}}$, all configurations with k pluses should appear. In other words, the threshold function for the appearance of a given pattern is $n^{-\frac{d}{2k}}$, where k is the number of +1's in the pattern. Proposition 4.4 below will show that the threshold function for a basic local sentence L is $n^{-\frac{d}{2k(L)}}$, where $k(L)$ is an integer that we call the *index* of L. Its definition refers to the decomposition (2.4) of a local property into a finite disjunction of complete descriptions, already used in the proof of Proposition 2.5.

Definition 4.3. *Let L be the basic local sentence defined by:*

$$\exists x_1 \ldots \exists x_m \left(\bigwedge_{1\leq i<j\leq m} dist(x_i,x_j) > 2r \right) \wedge \left(\bigwedge_{1\leq i\leq m} \psi_i(x_i) \right) .$$

If L is not satisfiable, then we shall set $k(L)=+\infty$. If L is satisfiable, for each $i=1,\ldots,m$, consider the finite set $\{D_{i,1},\ldots,D_{i,d_i}\}$ of those complete descriptions on the ball $B(x_i,r)$ which imply $\psi_i(x_i)$.

$$\psi_i(x_i) \leftrightarrow \bigvee_{1\leq j\leq d_i} D_{i,j}(x_i) .$$

Each complete description $D_{i,j}(x_i)$ corresponds to a configuration on $B(x_i,r)$. Denote by $k_{i,j}$ its number of $+1$'s.
The index *of L, denoted by $k(L)$ is defined by:*

$$k(L) = \max_{1\leq i\leq m} \min_{1\leq j\leq d_i} k_{i,j} . \tag{4.9}$$

The intuition behind Definition 4.3 is the following. Assume $e^{a(n)}$ is small compared to $n^{-\frac{d}{2k(L)}}$. Then there exists i such that none of the $D_{i,j}(x_i)$ can be satisfied, therefore there is no x such that $\psi_i(x)$ is satisfied, and L cannot be satisfied either. On the contrary, if $e^{a(n)}$ is large compared to $n^{-\frac{d}{2k(L)}}$, then for all $i=1,\ldots,m$, $\psi_i(x_i)$ should be satisfied for at least one vertex x_i, and the probability of satisfying L should be large. In other words, $n^{-\frac{d}{2k(L)}}$ is the threshold function of L.

Proposition 4.4. *Assume that for all n, $a(n)<0$ and there exists $b_0>0$ such that $|b(n)|<b_0$. Let L be a basic local property, and $k(L)$ be its index. If L is satisfiable and $k(L)>0$, then its threshold function for $e^{a(n)}$ is $n^{-\frac{d}{2k(L)}}$. If $k(L)=0$, its probability tends to 1.*

Proof: Assume L is satisfiable (otherwise its probability is null). For $a(n)<0$, we need to prove that $\mu_n(L)$ tends to 0 if $e^{2a(n)k(L)}n^d$ tends to 0 (in this case, $k(L)>0$), and that it tends to 1 if $e^{2a(n)k(L)}n^d$ tends to $+\infty$. The former will be proved first. Consider again the decomposition of L into complete descriptions:

$$L \leftrightarrow \exists x_1 \ldots \exists x_m \left(\bigwedge_{1\leq i<j\leq m} dist(x_i,x_j) > 2r \right) \wedge \left(\bigwedge_{1\leq i\leq m} \bigvee_{1\leq j\leq d_i} D_{i,j}(x_i) \right) .$$

If $e^{2a(n)k(L)}n^d$ tends to 0, there exists i such that:

$$\forall j=1,\ldots,d_i\,,\ \lim_{n\to\infty} e^{a(n)} n^{\frac{d}{2k_{i,j}}} = 0 .$$

By Proposition 4.2, the probability of $(\exists x\ D_{i,j}(x))$ tends to 0 for all $j=1,\ldots,d_i$. Therefore the probability of $(\exists x\, \psi_i(x))$ tends to 0, which implies that $\mu_n(L)$ tends to 0.
Conversely, for each $i=1,\ldots,m$, choose one of the $D_{i,j}$'s, such that the number of $+1$'s in the corresponding configuration is minimal (among all $k_{i,j}$'s). Denote that particular complete description by D_i and by k_i its number of $+1$'s (hence, $k(L)=\max_{1\leq i\leq m} k_i$). As in the proof of Proposition 4.2, we shall use the lattice T_n. Remember that its cardinality τ_n is of order n^d. The basic local sentence L is implied by the sentence L' defined as follows:

$$L' \leftrightarrow \exists x_1 \ldots \exists x_m \left(\bigwedge_{1\leq i\leq m} x_i \in T_n \right) \wedge \left(\bigwedge_{1\leq i<j\leq m} x_i \neq x_j \right) \wedge \left(\bigwedge_{1\leq i\leq m} D_i(x_i) \right) ,$$

For $i=1,\ldots,m$, let us introduce the random variable $\tilde{X}_n^{D_i}$:

$$\tilde{X}_n^{D_i} = \sum_{x\in T_n} I_x^{D_i} .$$

Even if some of the D_i's correspond to the same complete description, the event

$$\bigwedge_{i=1}^{m} \left\{ \tilde{X}_n^{D_i} \geq m \right\}$$

implies the sentence L' (and also the basic local sentence L). As a consequence, it is enough to prove the convergence of $\mu_n(\tilde{X}_n^{D_i} \geq m)$ to 1 for any index $i = 1, \ldots, m$.
Fix such an index i. We have already seen in the proof of Proposition 4.2 that, for two different vertices x and x' of T_n,

$$(i) \quad \mu_n \left(I_x^{D_i} = 1 \mid \mathcal{F}(\delta \mathcal{T}_n) \right) = \mu_n \left(I_{x'}^{D_i} = 1 \mid \mathcal{F}(\delta \mathcal{T}_n) \right) \; ;$$

$$(ii) \quad \mu_n \left(I_x^{D_i} = I_{x'}^{D_i} = 1 \mid \mathcal{F}(\delta \mathcal{T}_n) \right) = \mu_n \left(I_x^{D_i} = 1 \mid \mathcal{F}(\delta \mathcal{T}_n) \right) \times \mu_n \left(I_{x'}^{D_i} = 1 \mid \mathcal{F}(\delta \mathcal{T}_n) \right) \; .$$

These two statements mean that, conditioned on a given configuration σ on $\delta\mathcal{T}_n$, the distribution of the random variable $\tilde{X}_n^{D_i}$ is binomial with parameters τ_n and $\mu_n(I_x^{D_i} = 1 \,|\, \sigma)$. Moreover, for any vertex x, the inequality (4.6) says:

$$\exists\, c'' > 0, \; \forall\, \sigma \in \Xi_n, \; \tau_n \, \mu_n \left(I_x^{D_i} = 1 \mid \sigma_{\delta \mathcal{T}_n} \right) \geq c'' n^d e^{2a(n)k_i} \; . \tag{4.10}$$

Now, assume that $n^d e^{2a(n)k(L)}$ tends to infinity. Since $k_i \leq k(L)$ and $a(n)$ is negative, relation (4.10) involves the convergence of the product $\tau_n \times \mu_n \left(I_x^{D_i} = 1 \,|\, \sigma_{\delta \mathcal{T}_n} \right)$ to infinity, uniformly on $\sigma \in \Xi_n$. Then, it is easy to check that the conditional probability $\mu_n \left(\tilde{X}_n^{D_i} \geq m \,|\, \sigma_{\delta \mathcal{T}_n} \right)$ tends to 1, uniformly on the configuration $\sigma \in \Xi_n$. Therefore the probability

$$\mu_n \left(\tilde{X}_n^{D_i} \geq m \right) = \mathbb{E}_n \left[\mu_n \left(\tilde{X}_n^{D_i} \geq m \mid \mathcal{F}(\delta \mathcal{T}_n) \right) \right]$$

tends to 1 as n tends to infinity. □

Having characterized the threshold functions of all basic local properties, the proof of Theorem 4.1 is now clear. If $e^{a(n)} n^{\frac{d}{2k}}$ tends to 0 or $+\infty$ for any positive integer k, then by Proposition 4.4 the probability of any basic local sentence tends to 0 or 1. This remains true for any boolean combination of basic local sentences (cf. Proposition 2.5). By Gaifman's theorem, these boolean combinations cover all first-order sentences. Hence the zero-one law holds for first-order logic.

References

K.J. Compton. 0-1 laws in logic and combinatorics. In I. Rival, editor, *Algorithms and order*, pages 353–383. Kluwer, Dordrecht (1989).

D. Coupier. Poisson approximations for the Ising model. *J. Stat. Physics* (2006). To appear.

D. Coupier, A. Desolneux and B. Ycart. A zero-one law for first order logic on random images. In M. Drmota et al, editor, *Proc. Colloquium on Mathematics and computer science*, pages 495–505. Birkhäuser, Basel (2004).

R.L. Dobrushin and S.B. Shlosman. Completely analytical Gibbs fields. In *Statistical physics and dynamical systems (Köszeg, 1984)*, volume 10 of *Progr. Phys.*, pages 371–403. Birkhäuser Boston, Boston, MA (1985).

P. Doukhan. *Mixing: Properties and examples*, volume 85 of *L.N. in Statistics*. Springer-Verlag, Berlin (1994).

P. Doukhan and G. Lang. Rates of convergence in the weak invariance principle for the empirical repartition process of weakly dependent sequences. *Stat. Inf. for Stoch. Proc.* **5**, 199–228 (2002).

P. Doukhan and S. Louhichi. A new weak dependence condition and applications to moment inequalities. *Stoch. Proc. Appl.* **84**, 313–342 (1999).

P. Doukhan, G. Teyssière and P. Winant. Vector valued ARCH(∞) models. *To appear* (2006).

H.D. Ebbinghaus and J. Flum. *Finite model theory.* Springer-Verlag, Berlin (1995).

R. Fagin. Probabilities on finite models. *J. of Symbolic Logic* **41**, 50–58 (1976).

W. Feller. *An introduction to probability theory and its applications*, volume II. Wiley, London, 2nd edition (1971).

H. Gaifman. On local and non local properties. In J. Stern, editor, *Logic Colloquium '81*, pages 105–135. North Holland, Amsterdam (1982).

H.O. Georgii. *Gibbs measures and phase transitions*. de Gruyter, Berlin (1988).

Y.V. Glebskii, D.I. Kogan, M.I. Liogonkii and V.A. Talanov. Range and degree of realizability of formulas in the restricted predicate calculus. *Cybernetics* **5**, 142–154 (1969).

Y. Higuchi. Coexistence of infinite $(*)$-clusters. II. Ising percolation in two dimensions. *Probab. Theory Related Fields* **97** (1-2), 1–33 (1993).

R. Kindermann and J.L. Snell. *Markov random fields and their applications*, volume 1 of *Contemporary Mathematics*. AMS, Providence, RI (1980).

M. Rosenblatt. *Stationary processes and random fields*. Birkhäuser, Boston (1985).

J. Spencer. Zero-one laws with variable probability. *J. Symbolic Logic* **58**, 1–14 (1993).

J. Spencer. *The strange logic of random graphs*. Springer-Verlag, New York (2001).

M. Talagrand. On Russo's approximate zero-one law. *Ann. Probab.* **22**, 1576–1587 (1994).

P. Winkler. Random structures and zero-one laws. In Sauer, N.W. et al., editor, *Finite and infinite combinatorics in sets and logic*, pages 399–420. Kluwer, Dordrecht (1993).